Padmaka Dilukshan Mirihagalla

Evaluation, development, and application of organizational capability maturity models

Doctoral dissertation

SUMMARY

Supervisor: Professor Gyula Vastag

Győr

2022 September

Széchenyi István University

FOCUS OF THE THESIS

Enterprises are operating in increasingly dynamic business environments and sustaining competitive advantage requires continuously improving organizational capabilities (OCs). The fundamental idea of assessing OCs is to be able to understand the status quo, identify improvement opportunities, and develop to a higher level that leads to improved business performance. However, measurement approaches of OCs have been ill-equipped to objectively fulfill this mission. It is widely accepted that the developing OCs pass through stages of maturity and this evolution can be deliberated. Maturity models (MMs) emerged in this context and claimed to be designed to measure the maturity of OCs.

MMs are based on the premise that improved maturity in organizational capabilities leads to improvements in the desired outcomes. This promising potential explains the growing popularity of MMs in various academic journals. However, MMs have often been criticized for their lack of

methodological rigor and empirical validation. Despite such criticism, MMs have been a popular publication topic in journals of all ranking quartiles.

Given this context, this thesis project has produced four research papers resolving four critical issues built on each other. The first thoroughly explores the dynamics of MMs within the world of publications across dimensions such as journal rankings and citations. The second investigates theoretical issues within current MMs and provides the necessary theoretical grounding. The third demonstrates how to evaluate MMs using the introduced theoretical grounding in paper two. Finally, the fourth demonstrates through an empirical study how a robust MM should be built and applied within a chosen OC domain.

1. BACKGROUND

Enterprises increasingly operate in dynamically changing environments. To stay competitive, it is imperative that they develop organizational capabilities (OCs). OCs play

a central role in a multitude of business research domains (e.g., strategy, competitive advantage, learning, process management, people capabilities, organizational culture, and technology). Grant (1996, p. 377) defines an organizational capability as the "ability to perform repeatedly a productive task which relates either directly or indirectly to a firm's capacity for creating value through effecting the transformation of inputs into outputs". Collis (1994) claims that these OCs are 'socially complex routines' that help create defined outcomes within an enterprise. The underlying assumption of the capability point of view is that enterprises have ways of doing things and approaching problems that show elements of continuity (Dosi et al., 2003). Derivation of organizational capabilities occurs through its portfolio of skills, structures, processes, and norms supporting their integration (Teece et al., 1990). Through deliberate managerial decisions, enterprises can embed specific capabilities within their organizational fabric (Slotegraaf, 1999; Grewal & Slotegraaf, 2007).

Research on strategic management clearly exposes that the attention given to theory development has been excessive compared to empirical testing of them (Oxley et al., 2010; Bettis, 2012). Research on OCs has not been an exception to this, and criticism has specifically targeted the lack of empirical verification as opposed to the vast array of theory development (Newbert, 2007; Leiblein, 2011). The key issue and related problems with OCs are that they are latent constructs that are inherently unobservable (Grant & Verona, 2015). Furthermore, they are routinized, socially embedded, causally vague, and in particular, they don't necessarily reside only within a given function of an enterprise. Unfortunately, for OCs to fulfill their mission, most of the approaches used are ill-equipped in problem identification and especially objective measurement (Kruger & Dunning, 1999).

Most studies deploy self-reported questionnaires for measuring organizational capabilities – typically using the responses from the senior managers of responding enterprises (e.g., Snow and Hrebiniak, 1980; Kusunoki et

al., 1998; Sirmon et al., 2010). Issues with the accuracy and inconsistency of self-reported measurement approaches for OCs have been evident over decades of research (Kruger & Dunning, 1999). Deliberate misreporting, perceptual and cognitive distortion, and respondents not having visibility of the reality, are major issues in self-reported assessments in OC research (Huber & Power, 1985).

The fundamental idea behind assessing capabilities is to be able to assess the status quo for the enterprises and to be able to derive inferences for the scientific community (Nielsen & Kimberly, 1976). Assessment and improvement in OCs bring together assumptions from several scientific domains such as organizational change theory (Van de Ven & Poole, 1995), the resource-based view of a firm (Wernerfelt, 1984), life cycles, and teleological theories of goal formation (Lee & Kim, 2001). OCs develop, passing stages of maturity (Kohlegger et al., 2009) and over time, deliberate improvement attempts can help achieve consistency at higher levels of performance (Schumacher et al., 2016). The software development industry first

recognized this need and used it as justification to achieve higher degrees of reliability and validity in measuring OCs (Monteiro & Maciel, 2020). With the dawn of the 1990s, the software development industry was tired of unfulfilled promises of productivity and quality gains from OC measurement approaches with enterprises facing delays in delivery and excessive costs (Humphrey, 1988; Paulk et al., 1993).

> *"Even in undisciplined organizations, however, some individual software projects produce excellent results. When such projects succeed, it is generally through the heroic efforts of a dedicated team, rather than through repeating the proven methods of an organization with a mature software process. In the absence of an organization-wide software process, repeating results depends entirely on having the same individuals available for the next project. The success that rests solely on the availability of specific individuals provides no basis for long-term productivity and quality improvement throughout an organization. Continuous improvement*

can occur only through focused and sustained effort towards building a process infrastructure of effective software engineering and management practices." (Paulk et al., 1993, p. 31)

Within this context, the capability maturity model (CMM) (Paulk et al., 1993) was developed at the request of the US Department of Defense, Carnegie Mellon University. CMM defines the critical elements of a matured process and describes how an enterprise could move from lower maturity levels to higher maturity levels (Day-Lutteroth, 2011). Most importantly it provides clear descriptive definitions of each maturity level of the given capability (software development process management) resulting in improved reliability and validity in assessment compared to the approaches that existed before. Later, CMM was further improved into CMMI (CMM Integration) in 2001 with the integration of existing CMMs (Veldman-Klingenberg, 2009; Liou, 2011). By 2021, it is claimed that more than 10,000 businesses have adapted the CMMI method in over 106 countries (ISACA, 2021). As per many

systematic reviews, CMMI provided an influential methodological base and a turning point in the popularity of publications with MMs. Since then, the introduction of MMs from diverse capability domains has been apparent in academic research. MMs have been applied to capability domains such as healthcare, supply chain management, education, e-governance, project management, knowledge management, business process management, industry 4.0, lean manufacturing, product development, new product development, digital transformation, safety culture, sustainability, risk management, marketing, e-government, quality management, social media, blockchain, and the list keeps growing (Santos-Neto & Costa, 2019). MMs are descriptors that provide a stage-wise characterization of capabilities that organizations require to gauge their level of progress (Dekleva & Drehmer, 1997; Holland & Light, 2001; Fraser et al., 2002; Subba Rao et al., 2003; Gottschalk, 2009). Such hierarchical characterization is beneficial as MMs have the potential to aid performance excellence in many realms of human endeavor (Solli-Sæther & Gottschalk, 2010; van Looy et al., 2013). Pullen

(2007, p. 09) defines an MM as “a structured collection of elements that describe the characteristics of effective processes at different stages of development”.

2. ISSUES WITH MATURITY MODELS IN ACADEMIA

In my attempt to explore the domain of maturity approach to measuring OCs, the early part of the doctoral program was spent searching for and reading a significant number of publications that have introduced MMs. A special attempt was made to read as many publications as possible with MMs within the domains I had expert-level industry experience, e.g., Lean (e.g., Karlsson & Åhlström, 1996; Nightingale & Mize, 2002; Doolen & Hacker, 2005; Bhasin, 2011; Malmbrandt & Åhlström, 2013; Bijl et al., 2019; Zanon et al., 2020), project management (e.g., Kwak & Ibbs, 2002; Grant & Pennypacker 2006; Pasian et al., 2012), business process management (e.g, Tarhan et al., 2016; Van Looy et al., 2011) and quality management (e.g., Xiaofen, 2013; Ramadan & Arafeh, 2016). Once a deeper level of familiarity was achieved with the MMs of

these areas, attention was extended to understanding how the MM studies are constructed in relatively unfamiliar OC domains, e.g., industry 4.0 (e.g., Schumacher et al., 2016; Ganzarain & Nekane, 2016), risk management (Zhao et al., 2014; Salawu & Abdullah 2015), supply chain management (e.g., Lockamy & McCormack, 2004; Souza et al., 2015) and knowledge management (e.g., Serenko et al., 2016; Hsieh et al., 2009). Without attempting to be conclusive, this initial investigation surfaced some observations which later helped to frame the four research articles. Below are highlighted observations from the initial investigation of existing literature.

1. No MM other than CMMI has become widely accepted though two decades have passed since the introduction of CMMI.
2. Among the MMs within the domains where I had industry expertise, it was visible that structural components were missing and/or overlapping which made them unsuitable for making statistical inferences.

3. Some of the MMs had demonstrated how to use the model with an empirical example of application, however, only a few MMs had measured the capability maturity more than once. This was surprising as maturity is not a static concept and evolving maturity should be measured periodically.
4. There were no MMs with proven construct validity.
5. Almost no MM had been reused by other studies and most studies were only about introducing a new MM.
6. Published MMs come from all ranks of journals (Scimago) and many conference proceedings.
7. The objective of developing and publishing an MM seems to be academically biased as almost all the MMs justified the need for their model as "due to the absence of an MM within the capability domain" instead of the need for the MM to improve selected business performance indicators.

As maturity models have been a popular theme of publication over the last two decades, the number of reviews of them has also been common. There are many systematic reviews (e.g., Poeppelbuss et al., 2011;

Wendler, 2012; Tarhan et al., 2016; Reis et al., 2017; Vallerand et at., 2017; Santos-Neto & Costa, 2019), few recommended MM design processes (e.g., de Bruin et al., 2005; Becker et al., 2009) and a very few studies with design principles and evaluation criteria for MMs (Pöppelbuß & Röglinger, 2011; ISO, 2015). None of the systematic reviews contradicted my initial observations and most of these studies confirmed that current MMs are not developed as scientific measurement instruments based on the following two fundamentally important criteria: (1) MMs should be built on the 'predictable patterns' assumption that the higher maturity of capabilities leads to improved desired outcomes, and (2) they should have generalizability in achieving similar outcomes within proximal contexts.

3. THESIS STRUCTURE AND RESEARCH QUESTIONS

MMs are critical measurement instruments as they are associated with a range of benefits. These benefits include

gauging the status quo of capabilities (Becker et al., 2009), evaluating capabilities (Rosemann & de Bruin, 2005), supporting assessment (Iversen et al., 1999), enabling benchmarking (Domingues et al., 2016), guiding focus areas for investments in developing capabilities (Killen & Hunt, 2013; Vivares et al., 2018), normalizing skills deemed for improvement (Domingues et al., 2016), and demonstrating how a given approach has been evolving to achieve desired results (Santos-Neto & Costa, 2019). Apart from these industry benefits, they can be invaluable in research inquiries. In qualitative research, they can help structure investigations to understand deeper aspects of capabilities such as managing change, resistance to change, human-process synergies, collective wisdom, leadership orientation, etc. Provided they are built with rigor, MMs in quantitative research can establish a vast array of relationships between underexplored variables with capabilities such as levels of autonomy, empowerment, measurable leader behaviors, governance, business performance, employee satisfaction, client satisfaction, etc.

Most importantly, MMs provide a vital opportunity for understanding evolving relationships with such variables and capability maturity. Given such importance to both academia and industry, MMs have not lived up to their expectations. Often, MMs have been criticized for lacking theoretical rigor and empirical evidence of validity (Pöppelbuß & Röglinger, 2011; Santos-Neto and Costa, 2019; Monteiro & Maciel, 2020). Santos-Neto and Costa (2019) claim that out of journal-published MMs between 1973 and 2018, only 3% have been empirically validated and only 10% have demonstrated application. Wendler (2012) claims that MMs that had been useful in the real world are scarce, and many MMs have not even disclosed the methodological approach taken in their design. If MMs are intended to deliver performance excellence (van Looy et al., 2013), this situation is far from ideal for users in both academia and industry. Although with such a 'far-from-ideal' reality, MMs have been widely published in many low-to-high ranked peer-reviewed journals. This conflicting reality provided the direction to frame this

research project and exposed the four investigation opportunities:

1. Thoroughly understand the dynamics of MMs within the world of publications across dimensions such as journal rankings and citations.
2. Investigate theoretical issues within current MMs, provide theoretical grounding and establish a robust set of criteria to evaluate MMs.
3. Demonstrate how to evaluate MMs using the newly established criteria within a capability domain critical to both academia and the industry.
4. Demonstrate how a robust MM should be built and applied with real-life data through an empirical study.

These investigation opportunities were converted into separate research papers with multiple research questions being addressed. The four investigation opportunities are built on each other and represent critical yet different dimensions of the spectrum of MM research.

4. SUMMARIES OF PAPERS

4.1 Summary of research paper 1

Maturity models: Taking stock and moving forward

Coauthor: Professor Gyula Vastag
Journal: Hungarian Statistical Review (Classified as category A by Section 9 of the Hungarian Academy of Sciences)
Publication status: Published (June 2022 edition)

Maturity models (MMs) are based on the premise that improved maturity in organizational capabilities leads to improvements in the desired outcome measures. This promising potential explains the growing popularity of MMs and a large number of publications on the subject in various academic and professional journals. Numerous studies have conducted comprehensive systematic reviews providing a thematic analysis (e.g., Poeppelbuss & Simons, 2011; Reis et al., 2016; Wendler, 2012; Tarhan

et al., 2016; Vallerand et al., 2015; Santos-Neto & Costa, 2019; Monteiro & Maciel, 2020) and evaluations of MMs, along with their design principles (e.g., de Bruin et al., 2005; Becker et al., 2009; Pöppelbuß & Röglinger, 2011; ISO, 2015). Most systematic reviews involve a detailed descriptive statistical analysis of the themes of MMs and where they have been published (e.g., Santos-Neto & Costa, 2019; Monteiro & Maciel, 2020). However, no study has analyzed the qualitative parameters of journals in which this continuously growing number of MMs have been published. Given the increasing number of published studies on MMs, the broadening diversity of the capability domains of MMs, and the numerous criticisms of the methodological rigor of MMs, it is now interesting to explore the distribution of MMs across different journal ranks and to check whether publishing in higher-ranked journals means higher visibility or impact for MMs.

This review is based on the analysis of 339 papers published in 193 journals with MMs between 1973 and 2017 and answers the following RQs:

RQ1: How has the number of MMs published across Scimago Journal Ranking (SJR) quartiles changed over the years?

RQ2: How are the MMs published distributed across different journal categories?

RQ3: How is the distribution of MMs across journals and how are these journals ranked?

RQ4: Does publishing in higher-ranked journals result in more citations?

A publication's impact, measured by the number of cites collected over its lifetime, is influenced by the quality of the journal (measured by the journal's Article Influence Score by Clarivate Analytics, its Scimago Journal Ranking by Scimago, and the journal's Scimago quartile category), and the length of its availability to the public. Results have been derived from a variety of partitioning models (decision tree, bootstrap forest, boosted tree) and descriptive statistics. Although MMs have been heavily criticized for their lack of rigor, we see that they have been

published somewhat evenly across all the journal ranking levels. The very best top-ranked journals have not been interested in MMs and there might be a trend among the top journals to publish fewer MMs and this requires further investigation. Only 18% of the top 50% ranked journals have published MMs and only six journal categories account for the publication of 80% of the MMs indicating the potential to grow as a credible measurement instrument. There is a high chance of attracting more citations when published in higher-ranked journals and in journal rank deciles that have already published high volumes of MMs.

RQ1: *How has the number of MM articles published across SJR quartiles changed over the years?*

Since 2005, the number of published MM articles has increased; this finding is consistent with those of Wendler (2012), Santos-Neto & Costa (2019), and Monteiro and Maciel (2020). However, our analysis also shows that higher-ranked journals (Q1) do not immediately respond

to growth, increasing only in 2008. Since 2014, Q1 journals have been on a downward trend possibly because of the realization of the need for MMs as a scientific instrument requiring more rigor. Nevertheless, the reception of MMs within other quartiles is increasing. A new study covering the latest data is needed to test whether this trend has continued since 2018.

*RQ*2: *How are published MM articles distributed across different journal categories?*

Consistent with the findings of Santos-Neto and Costa (2019), our findings show that IT journals dominate in the publication of MM articles. Additionally, our analysis indicates that IT journals across all quartiles are receptive to MM articles. It will be interesting for future research to perform an in-depth analysis to determine whether the quality of MM articles published across different quartiles of IT journals differs from the set criteria. Engineering and operations/SCM journals form the second tier in terms of receptiveness towards MM articles. However, proportionately fewer Q4 journals within these two

categories have attracted MM articles. Capability domains within finance are a potential opportunity for MM articles to penetrate as currently only a very few finance journals have published MM articles (proportionate to the high volume of finance journals).

*RQ*3: *What is the distribution of MM articles across journals and how are these journals ranked?*

The only journal that publishes a high volume of MM articles is the *Software Quality Journal* (17 MM articles in total), and only six journals have published more than five MM articles. This indicates a high level of spread (339 MM articles across 193 journals) and implies that MM articles still have the potential to grow in number. Although the criticism over the rigor of MMs is intense, it is not visible in the analysis that only lower-ranked journals favor MM articles. Furthermore, the average net SJR (Scimago journal ranking) indicates that most of the journals increase their ranking during the period in which they publish MM articles.

Amongst the top 50% of business journals listed within Scimago, only 18% published MM articles from 1973 to 2017. Although the number of published MM articles has been on the rise since 2008, they have not yet evolved to become a fully accepted measurement instrument to gauge organizational capabilities. Different ranking groups (deciles within Q1 and Q2) have published different numbers of MM articles, and it appears that publishing within deciles with higher volumes of MM articles attracts a higher average number of citations. However, no MM article has been published in any of the top decile journals within Q1. This creates an opportunity to understand what it takes to publish an MM article in a journal that belongs to the top 10% of Q1 journals (the top 2.5% of all business journals).

*RQ*4: *Does publishing in higher-ranked journals result in more citations?*

Results from various partitioning models (regression tree, bootstrap forest, boosted tree) confirm what seems to have

been generally accepted from anecdotal evidence: publishing in top journals will have the greatest impact, and the AIS (Article influence score) measure of journal quality seems to be a good predictor.

4.2 Summary of research paper 2

Why are maturity models so juvenile? A critical review based on a new evaluation framework

Coauthor: Professor Gyula Vastag
Journal: Enterprise Information Systems (Scimago Q1, SJR 0.868)
Publication status: Submitted in Sep -2022
A version of this paper was previously rejected by,
The International Journal of Management Reviews (Scimago Q1, SJR 3.851) and Management Research Review (Scimago Q1, SJR 0.671)

This paper critically reviews fundamental theoretical issues with MMs, addresses them by providing in-depth

theoretical grounding, and introduces a new MM evaluation framework. The paper answers two RQs listed below.

RQ1: What are the theoretical issues with current MMs?
To answer this question, critical reviews, systematic reviews, and meta-analyses were explored creating a pool of existing criticism of MMs. Apart from these categories of studies, further criticism was searched for using keywords such as 'criticism', 'review', 'analysis', 'recommendations', and 'weaknesses' of maturity models to explore the current critical opinions of MMs. Both these searches were carried out in four digital databases (Web of Science, Scopus, SpringerLink, and ScienceDirect). The pool of criticism was cleaned-up to eliminate duplicates and items that were too broad to be able to thematically classify. The thematic analysis produced six broad categories of issues:

1. Clarity of the primary purpose of MMs
2. Clarity of the nature of the capabilities

3. Lack of studies with validated MMs
4. Not using an MM construction methodology/process
5. Missing specification or misspecification of the measurement model
6. Miscellaneous criticism

RQ2: What criteria should be used to evaluate MMS?

The objective of addressing this RQ is to provide a comprehensive framework with criteria that can be used to evaluate MMs. However, while answering RQ1, fundamental theoretical and conceptual issues of MMs were identified, and these had to be addressed first. The paper addresses these issues by providing theoretical grounding for MMs. Most criteria for evaluating MMs emerged from providing grounding for the six identified themes of criticism. However, to ensure that the framework is comprehensive, further criteria were searched for using keywords such as 'criteria for measurement models', 'formative index construction', 'measuring organizational constructs', and 'criteria in

evaluating maturity models'. The combined list of criteria was grouped based on common themes and specific criteria to be satisfied under each theme were identified. Answering RQ2, the themes and criteria were consolidated within a proposed framework.

The literature review explores the current schools of thought on MMs focusing on their claimed benefits, theoretical landscape, embedded assumptions, conceptual classifications, structural topology, and design processes. Then critical issues of MMs are exposed related to their primary purpose, the conceptualization of capabilities, validity, alignment of the design process, and specification of the measurement model. Addressing these issues, a firmer theoretical grounding for MMs is provided based on relevant multidisciplinary domains of research. A new MM evaluation framework is introduced consisting of OVRGP principles (opportunity, validity, reliability, generalizability, and process integrity), and their underlying criteria to be satisfied with the suggested

scientific tools. Figure 1 shows the new OVRGP conceptual framework.

Figure 1: OVRGP conceptual framework (maturity model evaluation criteria)

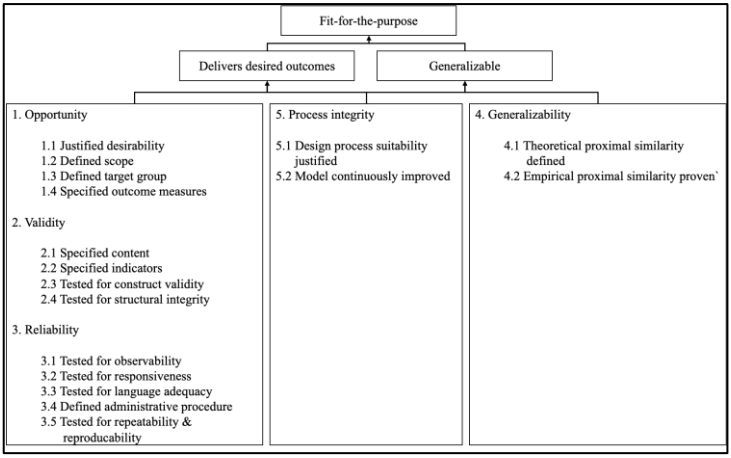


Table 1 lists the specific criteria that we propose an MM should meet under OVRGP principles. Where the term ‘expert opinion’ is mentioned, the use of grounded methods such as the Delphi technique, nominal group technique, case study interviews, and focus groups are

intended (De Bruin et al., 2005; Kimberlin & Winterstein, 2008).

Table 1: OVRGP principles and the underlying criteria

		Criteria descriptions
1. Opportunity	1.1	<u>Justified desirability</u> 1.1. a: How the measurement of maturity in the given domain (e.g., lean management) leads to business benefits should be justified.
	1.2	<u>Defined scope</u> 1.2.a: The scope of the domain should be clearly defined by stating its boundaries. 1.2.b: If similar domains exist, a differentiation should be made (e.g., lean vs. Total quality management).

	1.3	<u>Defined target group</u> 1.3.a: The target group, from which data is collected, should be clearly defined. 1.3.b: Qualifications of the target group should be stated. 1.3.c: Qualifications of the raters should be stated when self-assessment tools are suggested.
	1.4	<u>Specified outcome measures</u> 1.4.a: Stated business benefits in criteria 1.1, should be converted to measurable outcome variable(s) (e.g., lead time, waste reduction, inventory reduction in lean maturity). These defined outcome variables enable testing the model for construct validity (criteria 2.2 & 2.3).

2. Validity	2.1	<u>Specified content</u> 2.1.a: Domain components that represent the construct (e.g., lean management) should be specified. 2.1.b: Content within each domain component should be clearly defined. 2.1.c: The inclusion of each domain component should be justified based on their potential positive contribution to defined outcome measures in criteria 1.4.
	2.2	<u>Specified indicators</u> 2.2.a: The indicators specified should fully represent the specified domain components specified in criteria 2.1. 2.2.b: The inclusion of the specified indicators should be justified for their potential contribution to improving the outcome measures specified in criterion 1.4.

	2.3	<u>Tested for construct validity</u> 2.3.a: Theoretical construct validity should be achieved by defining the maturity stages in a clear path of meaningful hierarchical progression that leads to improvement in defined outcome variables in criterion 1.4. 2.3.b: Empirical construct validity should be proven with empirical data, showing that 'progressing along the stages (maturing) is positively correlated to improvements in the defined outcome variables in criterion 1.4.
	2.4	<u>Tested for structural integrity</u> 2.4.a: All the specified domain components and indicators within them should be tested for mutual exclusiveness and collective exhaustiveness (MECE).
4.	3.1	<u>Tested for observability</u> 3.1.a: Specified indicators should be defined with observable artifacts.

	3.2	Tested for responsiveness 3.2.a: The stage definitions should be precise in providing the rater with the ability to discriminate between the maturity levels.
	3.3	Tested for language adequacy 3.3.a: The language used in indicators and stage definitions should be understandable to the raters.
	3.4	Defined administrative procedure 3.4.a: User guidelines/administrative procedures on how the maturity model is to be used should be clearly outlined.
	3.5	Tested for repeatability & reproducibility (R&R) 3.5.a: The MM should be tested for R&R before full deployment. Gage R&R techniques used in the six sigma method or Kappa values are recommended.
4.	4.1	Theoretical proximal similarity defined 4.1.a: Proximal contexts should be stated where the MM is applicable in improving the specified outcome variables.

	4.2	<u>Empirical proximal similarity proven</u> 4.2.a Construct validity of the MM should be proven within a proximal context outside the context the construct validity has been proven in the first study.
5. Process integrity	5.1	<u>Design process suitability justified</u> 5.1.a: The design process of the MM used should be stated clearly and how the steps within the design process have ensured the criteria are met should be justified.
	5.2	<u>Model continuously improved</u> 5.2.a: Evidence should be provided on how the MM has been improved as a result of pre-testing or suggestions for improvement should be stated after the use in the empirical study.

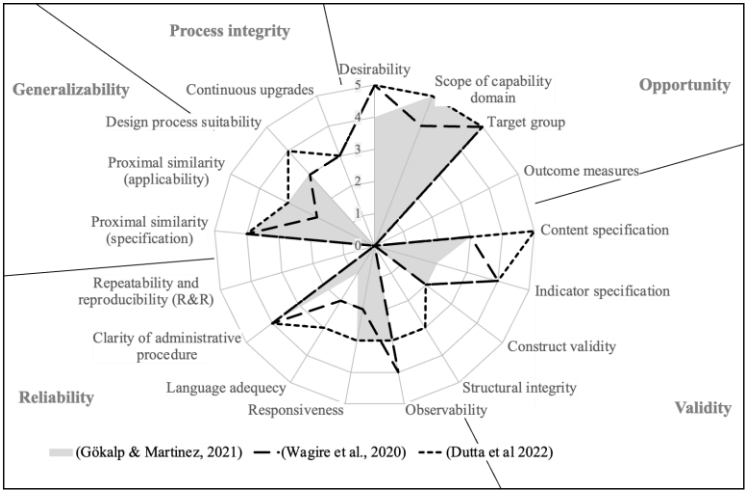
In order to demonstrate the use of the criteria, three selected MMs were evaluated, and the results are visualized in Figure 2. The assessment was performed using a five-point scale – (0) no evidence of the criteria

being considered, (1) criterion is mentioned but no evidence of fulfillment, (2) little evidence of criterion being fulfilled, (3) the criterion is partially fulfilled, (4), a major part of the criterion has been met, (5) the criterion is fully satisfied. This exercise is purely intended for illustration and visualization of where these MMs stand. Assessment is not based on the authors' judgment of the technical content of the MMs but based on the described methodology within the study. Below three MMs were chosen based on two criteria: (1) they are recent (assuming the rigor in MMs may have evolved), and (2) they are from Scimago quartile-1 journals (assuming strict peer review processes may have ensured rigor).

1) Digital transformation maturity assessment: development of the digital transformation capability maturity model (Gökalp & Martinez, 2021).
2) Development of maturity model for assessing the implementation of Industry 4.0: learning from theory and practice (Wagire et al., 2020).

3) An assimilation maturity model for IT governance and auditing (Dutta et al., 2022).

Figure 2: Criteria fulfillment radar chart for the chosen MMs.



The intention here is not to provide a full evaluation of the three MMs. However, we present here the most critical observations in summary form.

1. Opportunity: All three MMs are generally rated high; however, none have specified outcome measures.
2. Validity: The inclusion of a comprehensive literature review within the paper achieves content specification (exception: Dutta et al., 2022) and indicator specification is missing. All three MMs remain untested for construct validity using multiple rounds of empirical assessments.
3. Reliability: None of the three MMs show strong evidence of being pre-tested for observability, responsiveness, or language adequacy. There is no proof that any of the three MMs would have consistency in a multi-rater assessment.
4. Though generalizability has been claimed, they have been confined to the application only. There is no evidence of application in other entities that have resulted in statistically proven business improvements.

The evaluation framework can be used as a reference in structuring future systematic reviews, conducting meta-

analyses, evaluating MMs, gauging the thoroughness of new MMs in peer reviews, and most importantly, designing future MMs with built-in fit-for-purpose. We suggest studies proposing new MMs in the future use the proposed framework, show evidence of meeting the criteria, and where the criteria are not met, provide justifications. Capability maturity is a complex concept, and the act of operationalizing it through MMs encounters many potential failure points that are vulnerable to subjectivity. Fulfilling the proposed criteria might be viewed as a daunting research endeavor. However, given the abstract nature and the complexity of the capability maturity phenomenon, it is naive to expect currently established methods to produce fit-for-purpose MMs. Developing an MM and applying it in a single case study can give useful insights into the maturity phenomenon and we do not deny the qualitative usefulness of such insights. However, the ultimate gauge of a reusable scientific tool is proven only when it is empirically tested for construct validity, reliability, and generalizability. Developing an MM should not be the endgame. The idea behind MMs is

the premise that ‘better ways of working lead to better performance’. An MM can be criticized for its lack of rigor, however, for the concept of MM to be rejected, the said premise must be rejected. Therefore, the challenge is to keep testing and improving a given MM until it reaches an acceptable level of performance predictability. The essence of our paper is to convey this message to MM developers and to help them on the development path. We believe taking such a path will provide MMs the ability to understand the ‘capability maturity’ phenomenon deeper and gain much-needed industry acceptance by designing instruments with proven ability to improve performance.

4.3 Summary of research paper 3

Are the lean maturity models matured? A criteria-based evaluation

Coauthor: Professor Gyula Vastag
Journal: International Journal of Lean Six Sigma (Scimago Q1 in industrial and manufacturing, SJR 0.790)
Publication status: Submitted Sep-2022

Lean management is a journey of continuously maturing in the integration of processes, work systems, and capabilities within a convincing management philosophy. The journey of maturing in lean requires a compass to tell an enterprise its current position and the way ahead. An MM approach is intended to measure the degree of integration to lean management and the paper answers the following four RQs:

RQ1 - What criteria make an MM fit-for-the purpose?

RQ2 - How do the LMMs perform against these criteria?

RQ3 - What are the critical weaknesses of LMMs based on these criteria?

RQ4 - How can future LMMs be improved for their fit-for-purpose?

Note on RQ1: this RQ has been answered in paper 2. However, as the papers are independently submitted for publication, most of the content under this section is repeated.

5.3.1 Qualification of lean maturity models

The initial search was launched in four digital libraries – Web of Science, Scopus, SpringerLink, and ScienceDirect – using the following keywords: degree of leanness, lean maturity model, measuring lean adaptation, measuring lean integration, and measuring lean level. Four stages of filtering were performed sequentially to arrive at the sample used within the study. The sequence

of exclusion criteria was, (1) studies that are not from journals (e.g., conference papers, and books), (2) studies that do not include a measurement approach/model, (3) measurement models that have only used outcome measures, and (4) articles that have reused existing lean measurement models with no significant modifications. Since the focus of the study is closely related to assessing the theoretical rigor of existing LMMs, both conference papers and books were eliminated as such publications have not gone through a peer-review process. As explained in the literature review, not all measurement models use the term 'maturity', and terms such as leanness, and lean adaption level have been used as synonyms. Therefore, an additional round of filtering was conducted through in-depth scrutiny of the shortlisted models from step (4) above. The following questions were asked to perform this final layer of filtering based on the generic structure of a maturity model discussed within the literature review: (i) Does the model have a first-level breakdown of the construct similar to domain components? (ii) Does the model have indicators

representing the domain components identified in (i) above? And finally, (iii) are these indicators measured with a scale that reflects the degree of adaption? Irrespective of the terms used, if the answer was 'yes' to these questions, they were qualified to be within the sample of the study and are termed as LMMs. 27 LMMs were selected as the final list for this paper to be used for in-depth analysis and rating based on set criteria.

5.3.2 Establishing the theoretical framework for criteria

Due to the absence of a robust theoretical framework for designing MMs, a comprehensive screening of relevant literature was performed to arrive at the theoretical framework. The nature of MMs as measurement models was first explored before arriving at criteria as the criteria should have strong theoretical relevance. Relevant criteria were searched for using keywords such as 'criteria for measurement models', 'formative index construction', 'measuring organizational constructs', and 'criteria in evaluating maturity models'. The preliminary list of

identified criteria from the literature was grouped based on common themes and then the final list of criteria was selected based on relevance to MM's capability measurement models as well as formative models. We call the umbrella level identified themes the OVRGP criteria (opportunity, validity, reliability, generalizability, and process integrity) with 17 further sub-criteria (figure 1). Furthermore, the literature review has also been used to discuss the observations from the selected 27 LMMs against each criterion with examples of compliance as well as non-compliance.

5.3.3 Rating of lean maturity models based on the criteria

Each of the selected 27 papers and the embedded LMMs was studied in-depth to understand the degree to which each of the proposed 17 criteria has been met. In terms of preparation of qualitative data, all the evidence from each of the studies was tabulated with summarized extracts from the papers with 27 papers and 17 criteria on each axis. Once the qualitative extracts are revalidated for

accuracy, the rating was performed by using a 5-point scale; (0) no evidence of the criteria being considered, (1) criteria is mentioned but no evidence of fulfillment, (2) little evidence of criteria being fulfilled, (3) the criteria is partially fulfilled, (4) majority of the criteria have been met, (5) criteria are fully satisfied. The tabulated summaries were used by each author independently to rate the LMMs with inter-rater reliability calculated using Kappa values (Fleiss, 1971). When there was a conflict in rating, the original paper was referred to verify and arrive at a rating consensus between the authors and consequently achieving a high level of inter-rater reliability.

Figure 3 shows how the LMMs are ranked based on the 17 criteria within the OVRGP framework. The figure also shows the critical zones of weakness.

Figure 3: Sub-criteria in descending order based on the average rating (for all the LMMs).

	Karlsson and Ahlström (1996)	Nightingale and Mize (2002)	Soriano-Meier and Forrester (2002)	Sánchez and Pérez (2004)	Doolen and Hacker (2005)	Kollberg et. al., (2007)	Gurumurthy and Kodali (2009)	Singh et. al., (2010)	Bhasin (2011)	Vinodh and Chintha (2011)	Azevedo et. al., (2012)	Sezen et. al., (2012)	Malmbrandt and Åhlström (2013)	Kumar et. al., (2013)	Leyer & Moormann, (2014)	Maasouman and Demirli (2015)	Urban (2015)	Vidyadhar et. al., (2016)	Setianto and Haddud (2016)	Abdi (2018)	Santos Bento and Tontini (2018)	Yadav et. al., (2018)	Bijl et. al., (2019)	Galeazzo (2019)	Kaltenbrunner et. al., (2019)	Loyd et. al., (2020)	Zanon et. al., (2020)	
Desirability	5	4	3	2	2	2	3	2	3	2	1	3	5	3	4	3	3	5	5	4	3	3	3	3	2	4	4	
Scope	4	3	2	2	1	4	3	2	2	1	4	3	3	2	2	4	2	3	3	2	3	3	1	3	3	3	3	
Content specification	4	2	2	2	2	4	2	2	1	2	3	3	4	2	4	3	2	2	2	2	2	4	1	2	4	3	3	
Target group	3	4	4	3	3	1	4	2	0	1	1	4	4	1	2	1	2	2	3	3	3	1	4	3	4	4	1	
Clarity of administrative procedure	3	4	4	2	1	0	4	2	2	1	1	3	4	2	3	2	2	2	2	3	2	2	3	3	3	3	2	
Design process suitability	1	3	2	2	2	2	3	3	1	2	3	3	4	2	2	4	2	3	2	3	2	2	2	2	2	2	2	
Indicator specification	4	2	1	3	2	4	3	2	1	2	3	3	4	1	3	2	1	2	2	2	2	4	1	2	3	1	2	
Proximal similarity (specification)	3	3	3	3	3	3	2	2	2	2	2	2	3	1	3	2	0	2	1	4	1	2	2	2	2	3	2	
Outcome measures	4	2	0	3	1	2	3	0	0	0	0	4	4	0	0	2	0	0	0	0	5	0	5	5	5	5	0	Critical zone of weakness
Construct validity	2	2	0	1	1	1	2	0	1	2	2	3	2	1	2	2	1	0	3	0	3	2	4	4	4	4	1	
Responsiveness	0	2	1	1	1	1	0	1	2	2	1	1	3	1	1	4	2	2	0	2	1	1	3	4	1	2	3	
Observability	4	3	3	4	3	2	1	0	2	1	0	2	1	0	0	3	0	0	1	0	0	0	0	0	2	0	0	
Structural integrity	0	2	0	0	2	1	1	0	0	1	3	0	2	2	1	2	0	1	1	0	0	0	0	0	2	1	0	
Language adequacy	0	1	0	4	3	1	1	0	0	0	0	2	2	0	0	1	0	0	0	0	0	0	0	0	0	2	0	
Continuous upgrades	3	2	0	0	0	0	0	0	0	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	2	0	
Repeatability and reproducibility	0	1	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	1	1	0	1	1	
Proximal similarity (empirical)	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

The paper concludes that current LMMs are not built with methodological rigor that would enable them to be used as instruments to establish relationships between improving lean maturity and improving business performance. The evaluation reveals nine critical weaknesses of existing LMMs that include a lack of specified outcome measures, construct validity, rating scale responsiveness, indicator observability, structural integrity, language adequacy, evidence of continuous improvement, repeatability and reproducibility, and

empirical generalizability. For each critical weakness, remediations are recommended for future developers of LMMs. Noticeably current LMMs have performed relatively better in criteria that do not require empirical investigation or statistical testing. This gives the impression that the studies have taken a convenient approach to design and publishing instead of taking the effort to introduce LMMs with proven validity, reliability, and generalizability.

5.4 Summary of research paper 4:

Sustainable Entrepreneurship Orientation: A formative measurement model and application

Coauthor: Professor Gyula Vastag
Journal: Sustainability Accounting, Management and Policy Journal (Scimago Q1 in Business management and accounting, SJR 0.748)
Publication status: In review process. Revised version to be submitted by Nov-2022

Different versions of this paper had previously been rejected by the Journal of business ethics (Scimago Q1, SJR 2.438); Technological Forecasting and Social Change (after peer review) (Scimago Q1, SJR 2.336); Journal of cleaner production (Scimago Q1, SJR 1.921); Journal of Small Business Management (after peer review) (Scimago Q1, SJR 1.361); and Decision Sciences (Scimago Q1, SJR, 1.674).

The emergence of diverse Sustainable Entrepreneurship Orientation (SEO) perspectives has been evident in recent research. The introduction of a robust maturity measurement approach can accelerate the progress of SEO research through improved generalizability and comparability. This paper proposes an evidence-based approach to gauge the degree of SEO within an enterprise using a maturity model. The model consists of EPEC (Economic, People, Environment, Community) dimensions and their underlying observable indicators rated over pre-defined maturity levels. A comprehensive literature review identifies and explores the key

dimensions of the SEO domain. These dimensions are then integrated through the EPEC model and expanded into observable SEO firm-behavior indicators. The proposed SEO measurement model is based on the formative index construction rationale using multiple indicators and multiple causes (MIMIC) model. Using onsite observations and semi-structured interviews, quarterly rounds of maturing SEO levels were measured in a year-long investigation of a Hungarian private medical clinic to demonstrate the feasibility, and reliability of the approach. Fleiss' Kappa values ensured inter-rater reliability within this multi-rater exercise. The paper also discussed how the new measurement model will improve SEO research by offering new research possibilities, and its contribution to the industry as well as policymakers.

Figure 4 shows the new SEO conceptual model introduced based on the comprehensive literature review and analysis of content within each domain.

Figure 4: SEO conceptual model with dimensions and sub-dimensions.

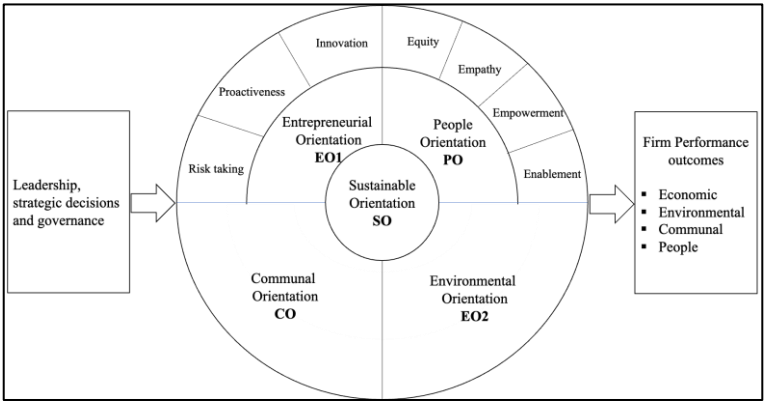


Table 2 shows the list of indicators derived from the literature within each of the dimensions identified in figure 4.

Table 2: Firm behavior indicators of SEO dimensions.

Π1 -EO1	*Π11:* Innovativeness	*x111*	Compared to the competition, the firm has introduced regular new products/services and has new products work in progress.
		x112	The firm regularly makes improvements to the existing products and services compared to the competition.
		x113	The firm constantly reviews existing processes and methods of operations and makes improvements.
		x114	The firm has dedicated human/financial resources allocated to develop new products/services and improve methods of operations.
	Π12: Risk Taking	*x121*	The firm commits to projects that can be considered high-risk for the size of the firm and its strength.
		x122	The firm commits a high portion of resources to opportunities that can be considered high-risk and high return.

		x123	During times of uncertainty, the firm takes bold decisions using uncertainty as an opportunity.
		x124	The firm experiments with new opportunities even if the outcomes are not fully certain.
		x125	Employees of the firm are encouraged and rewarded for bringing new ideas and taking risks.
	J13: Proactiveness	*x131*	Compared to the competition, the firm identifies and responds faster to opportunities.
		x132	The firm is usually the first to introduce new products and services to the market compared to the competition.
		x133	When competitors take new opportunities, the firm usually takes a bold non-defensive response.
		x134	The firm does not continue with dying products for long and takes action to discontinue them.

η2 -PO	*η21:* Empathy	*x211*	The firm makes a regular effort to gain insights into the emotional and cognitive concerns of employees.
		x212	The skills of leaders are continuously improved to understand and manage the emotional and cognitive concerns of employees.
		x213	There are formal and informal channels in place for employees to express their emotional and cognitive concerns.
		x214	Communication of change clearly takes the emotional and cognitive concerns of employees into consideration.
	η22: Equity	*x221*	The firm makes fair resource allocation decisions related to employees.
		x222	The design and implementation of firm policies related to employees are done with fairness as a consideration.
		x223	The management of the firm clearly communicates the rationale behind employee-related decisions.

	η_{23}: Empowerment	x_{231}	The firm regularly identifies and reduces practices that reduce employee work-related independence.
		x_{232}	Employees of the firm regularly receive participative opportunities to contribute to transformative activities.
		x_{233}	The firm identifies and removes barriers to access to top management for employees.
		x_{234}	The firm’s management provides clear directions, access to information, and resources that enable employees to work.
		x_{235}	Employees are regularly provided with a sense of purpose for their individual roles and the firm's direction.
	η_{24}: Enablement	x_{241}	The firm has methods in place for employees to acquire skills and knowledge.
		x_{242}	The firm provides methods and opportunities for employees to share knowledge with other employees.

		x243	Employees are rewarded and appreciated for acquiring new skills and knowledge with self-initiative.
		x244	Employees are given opportunities to use the new knowledge and skills they have acquired.
η3 - EO2	EO2	*x311*	The firm develops eco-friendly new products and services and improves the eco-friendliness of current products and services.
		x312	The firm has procedures in place to continuously reduce fossil fuel and water consumption.
		x313	The firm has methods in place to continuously reduce pollution and the use of hazardous materials.
		x314	The firm allocates resources to develop skills and capabilities needed to improve eco-friendliness.

η4 – CO	CO	*x411*	The firm develops products and services that contribute positively to the community.
		x412	The firm includes beneficiaries from the communities in the process of producing and distributing products and services.
		x413	The firm includes social activities within its commercial plans and activities.

Figure 5 shows the measured SEO maturity levels over 4 quarters within the selected case study enterprise with onsite observations and semi-structured interviews.

Figure 5: Egészség Kft, SEO maturity over four quarters for each indicator.

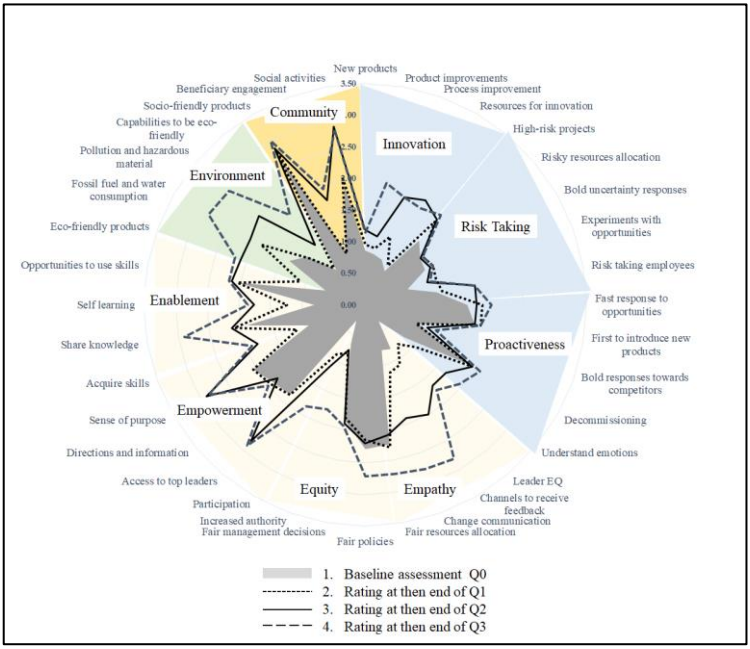
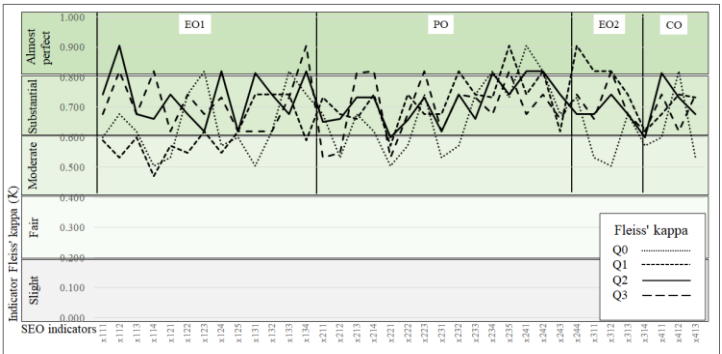


Figure 6 shows the proof of reliability among all the raters who participated.

Figure 6: Fleiss' Kappa values show high inter-rater reliability over four quarters.



	Baseline Q0	Q1	Q2	Q3
Overall Fleiss' kappa (K)	0.438	0.469	0.484	0.535
Standard Dev for all indicator Ks	0.113	0.104	0.090	0.075

This paper concludes by highlighting three vital outcomes: 1) the introduction of an SEO conceptual model with sub-component specifications, 2) the introduction of a formative measurement approach for SEO based on objective evidence-based firm behaviors, and 3) the demonstration of how the approach can be used in a combined qualitative-quantitative approach to investigate the SEO phenomena within a firm.

The proposed instrument and the method enable a wide array of avenues to credibly explore SEO within a firm, within multiple firms, or across industries. The approach makes it possible, quantitatively, to study factors influencing SEO, reflective relationships towards multi-dimensional firm performance caused by SEO using a MIMIC model; and qualitatively, to deeply understand the integration, evolution, change, and member behaviors within the SEO phenomena. Our case study gives a hint that the journey of becoming a matured SEO enterprise takes time, requires planned transformational efforts in multiple dimensions and the idea of SEO becomes more convincing to the employees over time. Apart from the academic contribution, we have also given user insights to practitioners, tech-platform developers, and policymakers. The SEO measurement approach includes observable behaviors written in a pragmatic language that firms can easily relate to and take necessary action to improve sustainable ways of working. In this way, we have also contributed to social science to improve its contribution to the environment and humanity.

CONTRIBUTION

This thesis is built around the fundamental premise that maturing in organizational capabilities leads to the desired performance increases. A thorough analysis of an extensive literature review revealed that many of the current maturity models are ill-equipped to represent the said premise. Despite these shortcomings, they are widely accepted and published in high-ranking scientific journals. This is an alarming reality for the science of organizational capabilities to ignore. On one hand, it stagnates the progress of organizational capability research, and on the other, it further distances academia from the industry. This thesis project does not claim to entirely solve this issue. However, using four research papers, the thesis project helps the science of organizational capabilities to move a few critical steps forward.

The first paper intends to be positioned as an eye-opener revealing that despite their significant weaknesses, maturity models have been published even in high-ranking

journals. Further, the paper shows that most MMs, irrespective of where they have been published, have gathered an insignificant number of citations given the significance of the OC domain. Though not among the primary objectives, the paper will also help developers of MMs to target their publication efforts in terms of journal selection, attracting citations, and selecting the right journal ranking parameters.

The second paper makes the most significant theoretical contribution. It reveals six fundamental theoretical issues related to MMs in their conceptual and construction rationale. The multidisciplinary theoretical grounding provided in this paper is expected to provide clearer direction for the future developers of MMs. The newly introduced OVRGP framework with 17 criteria will provide MMs with reliability, validity, and generalizability as scientific measurement instruments. Further, these 17 criteria will help future studies in structuring critical MM evaluations, and systematic reviews. Organizational capability research will progress

faster with MMs built based on these criteria and they will have much-needed industry acceptance.

The third paper puts the OVRGP framework and the 17 criteria into action. Lean management as an organizational capability is proven to deliver extraordinary performance improvements in many industries. This paper demonstrates to the lean management scientific community why the current MMs are ill-equipped to deliver their intended purpose. The paper will help future studies consider the OVRGP framework in developing LMMs that are reliable, valid, and generalizable. The detailed recommendations provided to overcome the identified nine critical weaknesses of LMMs further extend this contribution.

The fourth paper demonstrates how a maturity model should be built within the organizational capability of sustainable entrepreneurship. Apart from the demonstration of constructing an MM by combining

theory and empirical data, the paper makes a few important contributions to the scientific domain of sustainable entrepreneurship. Current measures are heavily biased towards outcomes and enterprise behavior-based measures are almost non-existent. The paper fills this void. Further, the sustainable entrepreneurship MM can help extend research in this relatively new domain. It has the potential for the industry users to establish a yardstick and the policymakers to establish enterprise behavior-based development programs and incentives.

6. LIST OF RESEARCH PROJECTS BY THE AUTHOR

Journal articles

1. Maturity models: Taking stock and moving forward, Journal: Hungarian Statistical Review (Published)
2. Why are maturity models juvenile? A critical review based on a new evaluation framework**,** Journal: Enterprise Information Systems (Submitted)

3. Are the lean maturity models matured? A criteria-based evaluation, Journal: International Journal of Lean Six Sigma (Submitted)
4. Sustainable Entrepreneurship Orientation: A formative measurement model and application, Journal: Sustainability Accounting, Management and Policy Journal (In review process)

Conference papers

1. Deploying lean transformation in services: An enterprise-level conceptual framework, Economic and social development conference, Hungary 2020 (Published)
2. Assessing the Impact-potential of AI on Finance Shared Services: A way forward, 1st IEEE international conference on hybrid corporate reality, Hungary 2022. (Submitted – Conference was cancelled)
3. Agent activation in agency-based insurance distribution: A case study, theoretical insights, and research opportunities, European Decision Sciences Institute Conference, Ireland 2022 (Published).

References

Becker, J., Knackstedt, R. and Pöppelbu, J. (2009) 'Developing Maturity Models for IT Management', *Business & Information Systems Engineering*, *1*(3), p213–222.

Bhasin, S. (2011) 'Measuring the Leanness of an organisation', *International Journal of Lean Six Sigma*, *2*(1), p55–74.

Bijl, A., Ahaus, K., Ruël, G., Gemmel, P. and Meijboom, B. (2019) 'Role of lean leadership in the lean maturity—second-order problem-solving relationship: a mixed methods study', *BMJ Open*, *9*(6), p026737.

de Bruin, T., Rosemann, M., Freeze, R. and Kulkarni, U. (2005) 'Understanding the main phases of developing a maturity assessment model', in *Proceedings of the Australasian Conference on Information Systems (ACIS)*, p8-19.

Domingues, P., Sampaio, P. and Arezes, P.M. (2016) 'Integrated management systems assessment: a maturity model proposal', *Journal of Cleaner Production*, *124*, p164–174.

Doolen, T.L. and Hacker, M.E. (2005) 'A review of lean assessment in organizations: an exploratory study of lean practices by electronics manufacturers', *Journal of Manufacturing Systems*, *24*(1), p55.

Dutta, A., Roy, R. and Seetharaman, P. (2022) 'An assimilation maturity model for IT governance and auditing', *Information & Management*, *59*, p1.

Fleiss, J.L. (1971) 'Measuring nominal scale agreement among many raters', *Psychological Bulletin*, *76*(5), p378–382.

Ganzarain, J. & Errasti, N. (2016) 'Three stage maturity model in SME's toward industry 4.0', *Journal of Industrial Engineering and Management*, *9*(5), p1119.

Gökalp, E. and Martinez, V. (2021) 'Digital transformation maturity assessment: development of

the digital transformation capability maturity model', *International Journal of Production Research*, p1-21.

Grant, K.P. and Pennypacker, J.S. (2006) 'Project management maturity: an assessment of project management capabilities among and between selected industries', *IEEE Transactions on engineering management*, *53*(1), p59-68.

Hsieh, P.J., Lin, B. and Lin, C. (2009) 'The construction and application of knowledge navigator model (KNM™): An evaluation of knowledge management maturity', *Expert Systems with Applications*, *36*(2), p4087–4100.

ISO. (2015) *ISO/IEC 33004: Information Technology – Process Assessment – Requirements for Process Reference, Process Assessment and Maturity Models*. International Standard Organization

Iversen, J., Nielsen, P.A. and Norbjerg, J. (1999) 'Situated assessment of problems in software

development', *ACM SIGMIS Database*, *30*(2), p66–81.

Karlsson, C. and Åhlström, P. (1996) 'Assessing changes towards lean production', *International Journal of Operations & Production Management*, *16*(2), p24–41.

Killen, C.P. and Hunt, R.A. (2013) 'Robust project portfolio management: capability evolution and maturity', *International Journal of Managing Projects in Business*, *6*(1), p131–151.

Kimberlin, C.L. and Winterstein, A.G. (2008) 'Validity and reliability of measurement instruments used in research', *American Journal of Health-System Pharmacy*, *65*(23), p2276–2284.

Kwak, Y.H. and Ibbs, C.W. (2002) 'Project Management Process Maturity (PM)2 Model', *Journal of Management in Engineering*, *18*(3), p150–155.

Lockamy, A. and McCormack, K. (2004) 'The development of a supply chain management process maturity model using the concepts of business

process orientation', *Supply Chain Management: An International Journal*, *9*(4), p272–278.

Monteiro, E.L. and Maciel, R.S.P. 2020 'Maturity Models Architecture: A large systematic mapping', *iSys-Brazilian Journal of Information Systems*, *13*(2), p110-140.

Nightingale, D.J. and Mize, J.H. (2002) 'Development of a lean enterprise transformation maturity model', *Information Knowledge Systems Management*, *3*, p15.

Pasian, B., Sankaran, S. and Boydell, S. (2012) 'Project management maturity: a critical analysis of existing and emergent factors', *International Journal of Managing Projects in Business*, *5*(1), p146–157.

Poeppelbuss, J. and Simons, A. (2011) 'Maturity Models in Information Systems Research: Literature Search and Analysis', *Communications of the Association for Information Systems*, *29*.

Pöppelbuß, J. and Röglinger, M. (2011) 'What makes a useful maturity model, A framework of general

design principles formaturitymodels and its demonstration in business processmanagement', in *Ecis Proceedings.*

Ramadan, N. and Arafeh, M. (2016) 'Healthcare quality maturity assessment model based on quality drivers', *International Journal of Health Care Quality Assurance*, *29*(3), p337–350.

Reis, T.L., Mathias, M.A.S. and de-Oliveira, O.J. (2016) 'Maturity models: identifying the state-of-the-art and the scientific gaps from a bibliometric study', *Scientometrics*, *110*(2), p643–672.

Rosemann, M. and de Bruin, T. (2005) 'Towards a business process management maturity model', in *ECIS 2005 Proceedings of the Thirteenth European Conference on Information Systems*, *1*, p1-12.

Salawu, R.A. and Abdullah, F. (2015) 'Assessing Risk Management Maturity of Construction Organisations on Infrastructural Project Delivery in Nigeria', *Procedia - Social and Behavioral Sciences*, *172*, p643–650.

Santos, G. and Tontini, G. (2018) 'Developing an instrument to measure lean manufacturing maturity and its relationship with operational performance', *Total Quality Management & Business Excellence*, *29*(9-10), p977–995.

Schumacher, A., Erol, S. and Sihn, W. (2016) 'A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises', *Procedia CIRP*, *52*, p161–166.

Serenko, A., Bontis, N. and Hull, E. (2016) 'An application of the knowledge management maturity model: the case of credit unions', *Knowledge Management Research and Practice*, *14*(3), p338-352.

Souza, R.P., Guerreiro, R. and Oliveira, M.P.V. (2015) 'Relationship between the maturity of supply chain process management and the organizational life cycle', *Business Process Management Journal*, *21*(3), p466–481.

Tarhan, A., Turetken, O. and Reijers, H.A. (2016) ‘Business process maturity models: A systematic literature review’ *Information and Software Technology*, *75*, p122–134.

Vallerand, J., Lapalme, J. and Moïse, A. (2015) ‘Analysing enterprise architecture maturity models: a learning perspective’, *Enterprise Information Systems*, *11*(6), p859–883.

Van Looy, A., Backer, M. and Poels, G. (2011) ‘Defining business process maturity. A journey towards excellence’, *Total Quality Management & Business Excellence*, *22*(11), p1119–1137.

Vivares, J.A., Sarache, W. and E. Hurtado, J. (2018) ‘A maturity assessment model for manufacturing systems’, *Journal of Manufacturing Technology Management*, *29*(5), p746–767.

Wagire, A.A., Joshi, R., Rathore, A.P.S. and Jain, R. (2020) ‘Development of maturity model for assessing the implementation of Industry 4.0:

learning from theory and practice', *Production Planning & Control*, p1–20.

Wendler, R. (2012) 'The maturity of maturity model research: A systematic mapping study', *Information and Software Technology*, *54*(12), p1317–1339.

Xiaofen, T. (2013) 'Investigation on quality management maturity of Shanghai enterprises', *The TQM Journal*, *25*(4), p417–430.

Zanon, L.G., Ulhoa, T.F. and Esposto, K.F. (2020) 'Performance measurement and lean maturity: congruence for improvement', *Production P1lanning & Control*, p1–15.

Zhou, Y. (2012) 'Towards Capability Maturity Model of E-Learning Process', *Intelligent Information Management*, *4*(04), p95.